46

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Trustees of the University of Pennsylvania
                Wilson, James M.
                Fisher, Krishna J.

    (ii) TITLE OF INVENTION: Method for Recombinant Adeno-Associated
          Virus-Directed Gene Therapy

   (iii) NUMBER OF SEQUENCES: 4

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Howson and Howson
        (B) STREET: Spring House Corporate Cntr, PO Box 457
        (C) CITY: Spring House
        (D) STATE: Pennsylvania
        (E) COUNTRY: USA
        (F) ZIP: 19477

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: WO
        (B) FILING DATE:
        (C) CLASSIFICATION:

   (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/708,188
        (B) FILING DATE: 06-SEP-1996

   (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/729,061
        (B) FILING DATE: 10-OCT-1996

  (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Kodroff, Cathy A.
        (B) REGISTRATION NUMBER: 33,980
        (C) REFERENCE/DOCKET NUMBER: GNVPN.019CIP2PCT

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: 215-540-9200
        (B) TELEFAX: 215-540-5818

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10398 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: cDNA

47

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GAATTCGCTA GCATCATCAA TAATATACCT TATTTTGGAT TGAAGCCAAT ATGATAATGA      60
GGGGGTGGAG TTTGTGACGT GGCGCGGGGC GTGGGAACGG GGCGGGTGAC GTAGTAGTGT     120
GGCGGAAGTG TGATGTTGCA AGTGTGGCGG AACACATGTA AGCGACGGAT GTGGCAAAAG     180
TGACGTTTTT GGTGTGCGCC GGTGTACACA GGAAGTGACA ATTTTCGCGC GGTTTTAGGC     240
GGATGTTGTA GTAAATTTGG GCGTAACCGA GTAAGATTTG GCCATTTTCG CGGGAAAACT     300
GAATAAGAGG AAGTGAAATC TGAATAATTT TGTGTTACTC ATAGCGCGTA ATATTTGTCT     360
AGGGAGATCT GCTGCGCGCT CGCTCGCTCA CTGAGGCCGC CCGGGCAAAG CCCGGGCGTC     420
GGGCGACCTT TGGTCGCCCG GCCTCAGTGA GCGAGCGAGC GCGCAGAGAG GGAGTGGCCA     480
ACTCCATCAC TAGGGGTTCC TTGTAGTTAA TGATTAACCC GCCATGCTAC TTATCTACAA     540
TTCGAGCTTG CATGCCTGCA GGTCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA     600
CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA     660
ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC CCACTTGGCA     720
GTACATCAAG TGTATCATAT GCCAAGTACG CCCCCTATTG ACGTCAATGA CGGTAAATGG     780
CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC     840
TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT     900
GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT CAATGGGAGT     960
TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC GTAACAACTC CGCCCCATTG    1020
ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA TAAGCAGAGC TCGTTTAGTG    1080
AACCGTCAGA TCGCCTGGAG ACGCCATCCA CGCTGTTTTG ACCTCCATAG AAGACACCGG    1140
GACCGATCCA GCCTCCGGAC TCTAGAGGAT CCGGTACTCG AGGAACTGAA AAACCAGAAA    1200
GTTAACTGGT AAGTTTAGTC TTTTTGTCTT TTATTTCAGG TCCCGGATCC GGTGGTGGTG    1260
CAAATCAAAG AACTGCTCCT CAGTGGATGT TGCCTTTACT TCTAGGCCTG TACGGAAGTG    1320
TTACTTCTGC TCTAAAAGCT GCGGAATTGT ACCCGCGGCC GCAATTCCCG GGGATCGAAA    1380
GAGCCTGCTA AAGCAAAAAA GAAGTCACCA TGTCGTTTAC TTTGACCAAC AAGAACGTGA    1440
TTTTCGTTGC CGGTCTGGGA GGCATTGGTC TGGACACCAG CAAGGAGCTG CTCAAGCGCG    1500
ATCCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCTGG CGTTACCCAA CTTAATCGCC    1560
TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC    1620
CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCTT TGCCTGGTTT CCGGCACCAG    1680
AAGCGGTGCC GGAAAGCTGG CTGGAGTGCG ATCTTCCTGA GGCCGATACT GTCGTCGTCC    1740
CCTCAAACTG GCAGATGCAC GGTTACGATG CGCCCATCTA CACCAACGTA ACCTATCCCA    1800
TTACGGTCAA TCCGCCGTTT GTTCCCACGG AGAATCCGAC GGGTTGTTAC TCGCTCACAT    1860
```

48

```
TTAATGTTGA TGAAAGCTGG CTACAGGAAG GCCAGACGCG AATTATTTTT GATGGCGTTA     1920

ACTCGGCGTT TCATCTGTGG TGCAACGGGC GCTGGGTCGG TTACGGCCAG GACAGTCGTT     1980

TGCCGTCTGA ATTTGACCTG AGCGCATTTT TACGCGCCGG AGAAACCGC CTCGCGGTGA      2040

TGGTGCTGCG TTGGAGTGAC GGCAGTTATC TGGAAGATCA GGATATGTGG CGGATGAGCG     2100

GCATTTTCCG TGACGTCTCG TTGCTGCATA AACCGACTAC ACAAATCAGC GATTTCCATG     2160

TTGCCACTCG CTTTAATGAT GATTTCAGCC GCGCTGTACT GGAGGCTGAA GTTCAGATGT     2220

GCGGCGAGTT GCGTGACTAC CTACGGGTAA CAGTTTCTTT ATGGCAGGGT GAAACGCAGG     2280

TCGCCAGCGG CACCGCGCCT TTCGGCGGTG AAATTATCGA TGAGCGTGGT GGTTATGCCG     2340

ATCGCGTCAC ACTACGTCTG AACGTCGAAA ACCCGAAACT GTGGAGCGCC GAAATCCCGA     2400

ATCTCTATCG TGCGGTGGTT GAACTGCACA CCGCCGACGG CACGCTGATT GAAGCAGAAG     2460

CCTGCGATGT CGGTTTCCGC GAGGTGCGGA TTGAAAATGG TCTGCTGCTG CTGAACGGCA     2520

AGCCGTTGCT GATTCGAGGC GTTAACCGTC ACGAGCATCA TCCTCTGCAT GGTCAGGTCA     2580

TGGATGAGCA GACGATGGTG CAGGATATCC TGCTGATGAA GCAGAACAAC TTTAACGCCG     2640

TGCGCTGTTC GCATTATCCG AACCATCCGC TGTGGTACAC GCTGTGCGAC CGCTACGGCC     2700

TGTATGTGGT GGATGAAGCC AATATTGAAA CCCACGGCAT GGTGCCAATG AATCGTCTGA     2760

CCGATGATCC GCGCTGGCTA CCGGCGATGA GCGAACGCGT AACGCGAATG GTGCAGCGCG     2820

ATCGTAATCA CCCGAGTGTG ATCATCTGGT CGCTGGGGAA TGAATCAGGC CACGGCGCTA     2880

ATCACGACGC GCTGTATCGC TGGATCAAAT CTGTCGATCC TTCCCGCCCG GTGCAGTATG     2940

AAGGCGGCGG AGCCGACACC ACGGCCACCG ATATTATTTG CCCGATGTAC GCGCGCGTGG     3000

ATGAAGACCA GCCCTTCCCG GCTGTGCCGA AATGGTCCAT CAAAAAATGG CTTTCGCTAC     3060

CTGGAGAGAC GCGCCCGCTG ATCCTTTGCG AATACGCCCA CGCGATGGGT AACAGTCTTG     3120

GCGGTTTCGC TAAATACTGG CAGGCGTTTC GTCAGTATCC CCGTTTACAG GGCGGCTTCG     3180

TCTGGGACTG GGTGGATCAG TCGCTGATTA AATATGATGA AAACGGCAAC CCGTGGTCGG     3240

CTTACGGCGG TGATTTTGGC GATACGCCGA ACGATCGCCA GTTCTGTATG AACGGTCTGG     3300

TCTTTGCCGA CCGCACGCCG CATCCAGCGC TGACGGAAGC AAAACACCAG CAGCAGTTTT     3360

TCCAGTTCCG TTTATCCGGG CAAACCATCG AAGTGACCAG CGAATACCTG TTCCGTCATA     3420

GCGATAACGA GCTCCTGCAC TGGATGGTGG CGCTGGATGG TAAGCCGCTG GCAAGCGGTG     3480

AAGTGCCTCT GGATGTCGCT CCACAAGGTA AACAGTTGAT TGAACTGCCT GAACTACCGC     3540

AGCCGGAGAG CGCCGGGCAA CTCTGGCTCA CAGTACGCGT AGTGCAACCG AACGCGACCG     3600

CATGGTCAGA AGCCGGGCAC ATCAGCGCCT GGCAGCAGTG GCGTCTGGCG GAAAACCTCA     3660

GTGTGACGCT CCCCGCCGCG TCCCACGCCA TCCCGCATCT GACCACCAGC GAAATGGATT     3720

TTTGCATCGA GCTGGGTAAT AAGCGTTGGC AATTTAACCG CCAGTCAGGC TTTCTTTCAC     3780
```

49

```
AGATGTGGAT  TGGCGATAAA  AAACAACTGC  TGACGCCGCT  GCGCGATCAG  TTCACCCGTG    3840

CACCGCTGGA  TAACGACATT  GGCGTAAGTG  AAGCGACCCG  CATTGACCCT  AACGCCTGGG    3900

TCGAACGCTG  GAAGGCGGCG  GGCCATTACC  AGGCCGAAGC  AGCGTTGTTG  CAGTGCACGG    3960

CAGATACACT  TGCTGATGCG  GTGCTGATTA  CGACCGCTCA  CGCGTGGCAG  CATCAGGGGA    4020

AAACCTTATT  TATCAGCCGG  AAAACCTACC  GGATTGATGG  TAGTGGTCAA  ATGGCGATTA    4080

CCGTTGATGT  TGAAGTGGCG  AGCGATACAC  CGCATCCGGC  GCGGATTGGC  CTGAACTGCC    4140

AGCTGGCGCA  GGTAGCAGAG  CGGGTAAACT  GGCTCGGATT  AGGGCCGCAA  GAAAACTATC    4200

CCGACCGCCT  TACTGCCGCC  TGTTTTGACC  GCTGGGATCT  GCCATTGTCA  GACATGTATA    4260

CCCCGTACGT  CTTCCCGAGC  GAAAACGGTC  TGCGCTGCGG  GACGCGCGAA  TTGAATTATG    4320

GCCCACACCA  GTGGCGCGGC  GACTTCCAGT  TCAACATCAG  CCGCTACAGT  CAACAGCAAC    4380

TGATGGAAAC  CAGCCATCGC  CATCTGCTGC  ACGCGGAAGA  AGGCACATGG  CTGAATATCG    4440

ACGGTTTCCA  TATGGGGATT  GGTGGCGACG  ACTCCTGGAG  CCCGTCAGTA  TCGGCGGAAT    4500

TACAGCTGAG  CGCCGGTCGC  TACCATTACC  AGTTGGTCTG  GTGTCAAAAA  TAATAATAAC    4560

CGGGCAGGCC  ATGTCTGCCC  GTATTTCGCG  TAAGGAAATC  CATTATGTAC  TATTTAAAAA    4620

ACACAAACTT  TTGGATGTTC  GGTTTATTCT  TTTTCTTTTA  CTTTTTTATC  ATGGGAGCCT    4680

ACTTCCCGTT  TTTCCCGATT  TGGCTACATG  ACATCAACCA  TATCAGCAAA  AGTGATACGG    4740

GTATTATTTT  TGCCGCTATT  TCTCTGTTCT  CGCTATTATT  CCAACCGCTG  TTTGGTCTGC    4800

TTTCTGACAA  ACTCGGCCTC  GACTCTAGGC  GGCCGCGGGG  ATCCAGACAT  GATAAGATAC    4860

ATTGATGAGT  TTGGACAAAC  CACAACTAGA  ATGCAGTGAA  AAAAATGCTT  TATTTGTGAA    4920

ATTTGTGATG  CTATTGCTTT  ATTTGTAACC  ATTATAAGCT  GCAATAAACA  AGTTAACAAC    4980

AACAATTGCA  TTCATTTTAT  GTTTCAGGTT  CAGGGGGAGG  TGTGGGAGGT  TTTTTCGGAT    5040

CCTCTAGAGT  CGAGTAGATA  AGTAGCATGG  CGGGTTAATC  ATTAACTACA  AGGAACCCCT    5100

AGTGATGGAG  TTGGCCACTC  CCTCTCTGCG  CGCTCGCTCG  CTCACTGAGG  CCGGGCGACC    5160

AAAGGTCGCC  CGACGCCCGG  GCTTTGCCCG  GGCGGCCTCA  GTGAGCGAGC  GAGCGCGCAG    5220

CAGATCTGGA  AGGTGCTGAG  GTACGATGAG  ACCCGCACCA  GGTGCAGACC  CTGCGAGTGT    5280

GGCGGTAAAC  ATATTAGGAA  CCAGCCTGTG  ATGCTGGATG  TGACCGAGGA  GCTGAGGCCC    5340

GATCACTTGG  TGCTGGCCTG  CACCCGCGCT  GAGTTTGGCT  CTAGCGATGA  AGATACAGAT    5400

TGAGGTACTG  AAATGTGTGG  GCGTGGCTTA  AGGGTGGGAA  AGAATATATA  AGGTGGGGGT    5460

CTTATGTAGT  TTTGTATCTG  TTTTGCAGCA  GCCGCCGCCG  CCATGAGCAC  CAACTCGTTT    5520

GATGGAAGCA  TTGTGAGCTC  ATATTTGACA  ACGCGCATGC  CCCCATGGGC  CGGGGTGCGT    5580

CAGAATGTGA  TGGGCTCCAG  CATTGATGGT  CGCCCCGTCC  TGCCCGCAAA  CTCTACTACC    5640

TTGACCTACG  AGACCGTGTC  TGGAACGCCG  TTGGAGACTG  CAGCCTCCGC  CGCCGCTTCA    5700
```
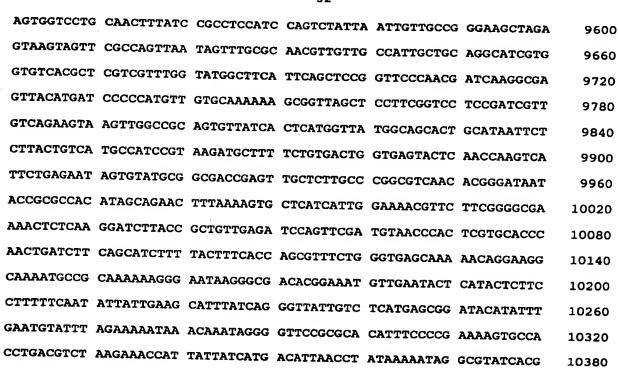
```
GCCGCTGCAG CCACCGCCCG CGGGATTGTG ACTGACTTTG CTTTCCTGAG CCCGCTTGCA   5760

AGCAGTGCAG CTTCCCGTTC ATCCGCCCGC GATGACAAGT TGACGGCTCT TTTGGCACAA   5820

TTGGATTCTT TGACCCGGGA ACTTAATGTC GTTTCTCAGC AGCTGTTGGA TCTGCGCCAG   5880

CAGGTTTCTG CCCTGAAGGC TTCCTCCCCT CCCAATGCGG TTTAAAACAT AAATAAAAAA   5940

CCAGACTCTG TTTGGATTTG GATCAAGCAA GTGTCTTGCT GTCTTTATTT AGGGGTTTTG   6000

CGCGCGCGGT AGGCCCGGGA CCAGCGGTCT CGGTCGTTGA GGGTCCTGTG TATTTTTTCC   6060

AGGACGTGGT AAAGGTGACT CTGGATGTTC AGATACATGG GCATAAGCCC GTCTCTGGGG   6120

TGGAGGTAGC ACCACTGCAG AGCTTCATGC TGCGGGGTGG TGTTGTAGAT GATCCAGTCG   6180

TAGCAGGAGC GCTGGGCGTG GTGCCTAAAA ATGTCTTTCA GTAGCAAGCT GATTGCCAGG   6240

GGCAGGCCCT TGGTGTAAGT GTTTACAAAG CGGTTAAGCT GGGATGGGTG CATACGTGGG   6300

GATATGAGAT GCATCTTGGA CTGTATTTTT AGGTTGGCTA TGTTCCCAGC CATATCCCTC   6360

CGGGGATTCA TGTTGTGCAG AACCACCAGC ACAGTGTATC CGGTGCACTT GGGAAATTTG   6420

TCATGTAGCT TAGAAGGAAA TGCGTGGAAG AACTTGGAGA CGCCCTTGTG ACCTCCAAGA   6480

TTTTCCATGC ATTCGTCCAT AATGATGGCA ATGGGCCCAC GGGCGGCGGC CTGGGCGAAG   6540

ATATTTCTGG GATCACTAAC GTCATAGTTG TGTTCCAGGA TGAGATCGTC ATAGGCCATT   6600

TTTACAAAGC GCGGGCGGAG GGTGCCAGAC TGCGGTATAA TGGTTCCATC CGGCCCAGGG   6660

GCGTAGTTAC CCTCACAGAT TTGCATTTCC CACGCTTTGA GTTCAGATGG GGGGATCATG   6720

TCTACCTGCG GGGCGATGAA GAAAACGGTT TCCGGGGTAG GGGAGATCAG CTGGGAAGAA   6780

AGCAGGTTCC TGAGCAGCTG CGACTTACCG CAGCCGGTGG GCCCGTAAAT CACACCTATT   6840

ACCGGGTGCA ACTGGTAGTT AAGAGAGCTG CAGCTGCCGT CATCCCTGAG CAGGGGGGCC   6900

ACTTCGTTAA GCATGTCCCT GACTCGCATG TTTTCCCTGA CCAAATCCGC CAGAAGGCGC   6960

TCGCCGCCCA GCGATAGCAG TTCTTGCAAG GAAGCAAAGT TTTTCAACGG TTTGAGACCG   7020

TCCGCCGTAG GCATGCTTTT GAGCGTTTGA CCAAGCAGTT CCAGGCGGTC CCACAGCTCG   7080

GTCACCTGCT CTACGGCATC TCGATCCAGC ATATCTCCTC GTTTCGCGGG TTGGGGCGGC   7140

TTTCGCTGTA CGGCAGTAGT CGGTGCTCGT CCAGACGGGC CAGGGTCATG TCTTTCCACG   7200

GGCGCAGGGT CCTCGTCAGC GTAGTCTGGG TCACGGTGAA GGGGTGCGCT CCGGGCTGCG   7260

CGCTGGCCAG GGTGCGCTTG AGGCTGGTCC TGCTGGTGCT GAAGCGCTGC CGGTCTTCGC   7320

CCTGCGCGTC GGCCAGGTAG CATTTGACCA TGGTGTCATA GTCCAGCCCC TCCGCGGCGT   7380

GGCCCTTGGC GCGCAGCTTG CCCTTGGAGG AGGCGCCGCA CGAGGGCAG TGCAGACTTT   7440

TGAGGGCGTA GAGCTTGGGC GCGAGAAATA CCGATTCCGG GGAGTAGGCA TCCGCGCCGC   7500

AGGCCCCGCA GACGGTCTCG CATTCCACGA GCCAGGTGAG CTCTGGCCGT TCGGGGTCAA   7560

AAACCAGGTT TCCCCCATGC TTTTTGATGC GTTTCTTACC TCTGGTTTCC ATGAGCCGGT   7620
```

51

```
GTCCACGCTC GGTGACGAAA AGGCTGTCCG TGTCCCCGTA TACAGACTTG AGAGGCCTGT    7680

CCTCGACCGA TGCCCTTGAG AGCCTTCAAC CCAGTCAGCT CCTTCCGGTG GGCGCGGGGC    7740

ATGACTATCG TCGCCGCACT TATGACTGTC TTCTTTATCA TGCAACTCGT AGGACAGGTG    7800

CCGGCAGCGC TCTGGGTCAT TTTCGGCGAG GACCGCTTTC GCTGGAGCGC GACGATGATC    7860

GGCCTGTCGC TTGCGGTATT CGGAATCTTG CACGCCCTCG CTCAAGCCTT CGTCACTGGT    7920

CCCGCCACCA AACGTTTCGG CGAGAAGCAG GCCATTATCG CCGGCATGGC GGCCGACGCG    7980

CTGGGCTACG TCTTGCTGGC GTTCGCGACG CGAGGCTGGA TGGCCTTCCC CATTATGATT    8040

CTTCTCGCTT CCGGCGGCAT CGGGATGCCC GCGTTGCAGG CCATGCTGTC CAGGCAGGTA    8100

GATGACGACC ATCAGGGACA GCTTCAAGGA TCGCTCGCGG CTCTTACCAG CCTAACTTCG    8160

ATCACTGGAC CGCTGATCGT CACGGCGATT TATGCCGCCT CGGCGAGCAC ATGGAACGGG    8220

TTGGCATGGA TTGTAGGCGC CGCCCTATAC CTTGTCTGCC TCCCCGCGTT GCGTCGCGGT    8280

GCATGGAGCC GGGCCACCTC GACCTGAATG GAAGCCGGCG GCACCTCGCT AACGGATTCA    8340

CCACTCCAAG AATTGGAGCC AATCAATTCT TGCGGAGAAC TGTGAATGCG CAAACCAACC    8400

CTTGGCAGAA CATATCCATC GCGTCCGCCA TCTCCAGCAG CCGCACGCGG CGCATCTCGG    8460

GCAGCGTTGG GTCCTGGCCA CGGGTGCGCA TGATCGTGCT CCTGTCGTTG AGGACCCGGC    8520

TAGGCTGGCG GGGTTGCCTT ACTGGTTAGC AGAATGAATC ACCGATACGC GAGCGAACGT    8580

GAAGCGACTG CTGCTGCAAA ACGTCTGCGA CCTGAGCAAC AACATGAATG GTCTTCGGTT    8640

TCCGTGTTTC GTAAAGTCTG GAAACGCGGA AGTCAGCGCC CTGCACCATT ATGTTCCGGA    8700

TCTGCATCGC AGGATGCTGC TGGCTACCCT GTGGAACACC TACATCTGTA TTAACGAAGC    8760

CTTTCTCAAT GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG    8820

GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT    8880

CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG    8940

ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC    9000

GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA    9060

AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT    9120

GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT    9180

TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA    9240

TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT TAAATCAATC    9300

TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT    9360

ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT GACTCCCCGT CGTGTAGATA    9420

ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA    9480

CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA    9540
```

52

```
AGTGGTCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA ATTGTTGCCG GGAAGCTAGA    9600

GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG CCATTGCTGC AGGCATCGTG    9660

GTGTCACGCT CGTCGTTTGG TATGGCTTCA TTCAGCTCCG GTTCCCAACG ATCAAGGCGA    9720

GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT    9780

GTCAGAAGTA AGTTGGCCGC AGTGTTATCA CTCATGGTTA TGGCAGCACT GCATAATTCT    9840

CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG GTGAGTACTC AACCAAGTCA    9900

TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC CGGCGTCAAC ACGGGATAAT    9960

ACCGCGCCAC ATAGCAGAAC TTTAAAAGTG CTCATCATTG GAAAACGTTC TTCGGGGCGA   10020

AAACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA TGTAACCCAC TCGTGCACCC   10080

AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG GGTGAGCAAA AACAGGAAGG   10140

CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGGAAAT GTTGAATACT CATACTCTTC   10200

CTTTTTCAAT ATTATTGAAG CATTTATCAG GGTTATTGTC TCATGAGCGG ATACATATTT   10260

GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA CATTTCCCCG AAAAGTGCCA   10320

CCTGACGTCT AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG   10380

AGGCCCTTTC GTCTTCAA                                                 10398
```
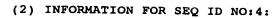
(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
CATGGTAATA GCGATGACTA                                                   20
```

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GCTCTGCTTA TATAGACCTC                                                   20
```

53

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATAAGCTGCA ATAAACAAGT   20